Analytic Provenance for Sensemaking: A Research Agenda

Journal:	*IEEE Computer Graphics and Applications*
Manuscript ID:	Draft
Manuscript Type:	Special Issue - VIS Workshop
Date Submitted by the Author:	n/a
Complete List of Authors:	Xu, Kai; Middlesex University, attfield, simon; Middlesex University, Jankun-Kelly, T.J.; Mississippi State University, Department of Computer Science and Engineering Wheat, Ashley; Middlesex University, Nguyen, Phong; Middlesex University, Selvaraj, Nallini; Middlesex University,
Keywords:	I.6.9.c Information visualization < I.6.9 Visualization < I.6 Simulation, Modeling, and Visualization < I Computing Methodologies, H.1.2.b Human-centered computing < H.1.2 User/Machine Systems < H.1 Models and Principles < H Information Technology and Systems, H.5.2.f Graphical user interfaces < H.5.2 User Interfaces < H.5 Information Interfaces and Representation (HCI) < H Information Technology and Systems

Analytic Provenance for Sensemaking: A Research Agenda

Kai Xu Simon Attfield T.J. Jankun-Kelly
Ashley Wheat Phong H. Nguyen Nallini Selvaraj

Sensemaking is a process of finding meaning from information — a process of comprehension. A common theme is the idea of sensemaking as the construction, elaboration and reconciliation of representations which account for and explain the information we receive about the world. Things make sense when we have a representation or generalized belief which fits to our experience, and things fail to make sense when this coherence is missing or there is simply no representation to speak of. Important to the sensemaking dynamic is the reciprocal relationship between the two, i.e. representations create expectations which guide us in information that we seek and also how we interpret information when we receive it, whilst information can challenge and shapes the representations we create.

During complex sensemaking tasks, it can be valuable to maintain a history of the data and reasoning involved – referred to as *provenance* information. Provenance information can be a resource for "reflection-in-action" during analysis, supporting collaboration between analysts, and reporting to decision makers. It can also act as a resource after the event, supporting the interpretation of claims, audit, accountability, and training.

There has been considerable work on capturing and visualizing *data provenance*, which focuses on data collection and computation, and *analytic provenance*, which captures the interactive data exploration and reasoning process. However, there is limited work of utilizing such provenance information to support sensemaking, in terms of improving efficacy and avoiding pitfalls such as issues of data quality and human bias. A workshop was held during IEEE VIS 2014 with the aim of bringing together researchers involved in visual analytics and various aspects of sensemaking to bridge this gap. The workshop participants considered emerging positions and findings related to the capture, processing, representation and use of provenance information to support complex sensemaking tasks. In this article, we present and extend the research challenges discussed in the workshop in order to provide a agenda for sensemaking analytic provenance.

1 Modelling Analytic Provenance

Analytic provenance information can be categorized into a hierarchical model based on its semantic richness [5]. An example of analyzing the stock market (i.e., the sensemaking task) is shown in Figure 1: the level of semantics in the analytic provenance increases from bottom to top. The bottom-level *events* consists of low-level user interactions such as mouse-clicks and keystrokes, which have little semantic meaning. The next level up is *actions*. These are analytic steps such as querying the database or changing the zooming level of data visualization. The data and visualization involved in such actions are also part of the analytic provenance at this level. Further up are the *sub-tasks*, which in many cases are the analyses required to achieve the sensemaking goal. In the context of stock market analysis, examples are identifying top performing companies and determining long term trends. In the top-level is the *task*, which is "analyzing stock market".

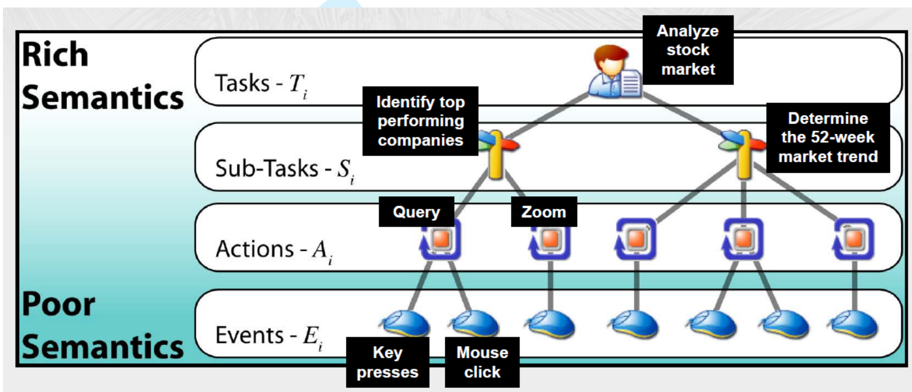

Figure 1: The hierarchical analytic provenance model shown with an example of analyzing stock market. The semantic richness increases from bottom to top. The bottom layer are the *events* such as key presses and mouse clicks, which have little semantics. The next level up are *actions* such as the database query and visualization zooming. Further up are the *sub-tasks*, which usually are the analyses performed during the sensemaking. The top level *tasks* are the sensemaking goals.

Analytic provenance information is closely linked both within and across layers. Within a layer, analytic provenance is linked temporally (i.e. one event happens after another) and logically (i.e. one action depends on the two previous actions). The connections are also cross layers: a database query action can consist of several mouse click and key stroke events; several analysis actions are usually required to finish an analysis at the sub-task level.

The model can be considered an abstraction hierarchy of goals or tasks. From any point in the hierarchy, looking up gives you the 'why' and looking down gives you the 'how'. The 'why' also provides explanatory context context or rationale for the 'how'. Task performance moves from right to left at all levels simultaneously. Outcomes can also be described at any level: A low-level outcome that a bar in a bar chart is taller than the rest corresponds to a higher-level result of relative stock performance. We assume the the intent of the process as a whole it too high level to be adequately described by any individual component of the model, not without some knowledge based interpretation: Analyzing the stock market for an investment opportunity would be such an example. Although this goal might be inferred from the observation, the inference depends as much on knowledge of people and the kinds of things they might do when they want to find an investment opportunity as it is on the observation that they are analyzing the stock market.

The hierarchical structure of the analytic provenance has considerable implications for its capture and visualization. It is equally important to capture and visualize individual pieces of analytic provenance information and the connections between them. Communicating the provenance during collaboration while insuring appropriate user privacy is also a challenge. Finally, detecting, communicating, and dealing with uncertainty in the the hierarchical process—building trust in the final result—is also vital. These were the primary facets of our workshop discussion and will be presented here in more depth.

2 High-level Analytic Provenance Capture

Analytic provenance capture provides the data for further analysis and visualization. What is captured and its quality decides what analysis and visualization are possible and their quality.

Capturing lower level events and actions is relatively straightforward in a visual analytics system. However, such analytic provenance information alone is of limited use [5]. Tasks and sub-task information provide important clues to the purpose and rationale behind the events and actions a user performed during sensemaking. However, they are largely part of users' thinking, which a visual analytics system does not have direct access to. This is one of the biggest challenges in analytic provenance capture. There is a limited time window to capture such information, because after a while even the users themselves may forget what they were doing, at which point it becomes very difficulty to recover the analytic provenance information.

Existing approaches to capture high level analytic provenance can be broadly categorized into *manual* and *automatic* methods. The manual methods largely rely on users recording the details of their analysis and sensemaking tasks, whereas the automatic methods try to infer the higher level analytic provenance from lower level events and actions that are automatically captured. While the manual approaches are usually more accurate, it can distract user from the actual analysis task and such extra cognitive effort may discourage users from recording analytic provenance. On the other hand, the automatic approaches do not introduce interruption to the sensemaking process, but their capability of inferring semantic-rich analytic provenance information is limited [5]. Personal differences introduces additional difficulty for automatically inferring high level analytic provenance. Users' knowledge and experience have a considerable impact on the way they conduct sensemaking analysis. As a result, the sensemaking process (i.e. the analytic provenance) can vary significantly from user to user, even with the same dataset and analysis task.

2.1 Manual Capture

Since it is possible to automatically capture analytic provenance at event and action level, manual capture mostly focuses on the task and sub-task level. Allowing user annoation is one of the most common forms: User creates *notes* that are associated with certain data, analysis result, or visualization. A "note" is not limited to record interesting data patterns; it can also describe the causal relationships between patterns and even the reasoning process that generates hypothesis based on existing findings. *Data-aware* annotation links the visualization and analysis results to the underlying data used to produce them, which makes it possible to apply new analysis and visual mapping at a later stage if further investigation is needed.

While individual note only represents a fraction of the analytic provenance, it is possible to provide a reasonably good overview of the sensemaking process if a number of notes and the connections between them are captured. However, this is only possible when users are willing to take notes, which can be perceived as distractions sometimes. There are two common strategies to alleviate this:

- Minimize interruption and cognitive effort;
- Provide tangible benefits to the sensemaking task.

Reducing interruption and cognitive effort can lower the likelihood that users are discouraged from recording analytic provenance. This can be achieved through integration with the analysis tools (so users do not need to switch between interfaces) or streamlining the recording process (e.g., with minimal mouse clicking and movement). Besides, it is likely to motivate user adoption if the analytic provenance captured can provide perceivable benefits to the analysis task, i.e., immediate support of sensemaking process. Examples include the ability to record discoveries during the analysis and review and plan exploratory analysis for complex sensemaking task. However, currently there is a lack of general design guidelines for how to achieve them, and there are few user studies evaluating how effective they are. Any progress related to these two challenges can have a considerable impact on the capture of analytic provenance and enable better support for the sensemaking.

2.2 Automatic Capture

One of the main disadvantages of manual capture is the requirement of direct input from the users. Automatic approaches try to address this by inferring missing analytic provenance information from those that can be automatically captured. As discussed earlier, it is easier to capture analytic provenance at the event and action level. Therefore, most automatic approaches try to infer sub-task- and task-level information from event and action analytic provenance.

This turns out to be a difficult challenge. An experiment studied how much of a user's reasoning process can be recovered from user action information [2]. A domain-specific sensemaking task was used and experts were recruited to analyze the user action log. Higher-level analytic provenance manually inferred from the interaction logs were compared with the ground truth obtained through interview. The results showed that 79 percent of the findings, 60 percent of the methods, and 60 percent of the strategies were correctly recovered. The accuracy is not high even in such a constrained setting with domain experts doing the inference. Given the diversity of data and analysis involved in the sensemaking and the difficult of replicating expert knowledge/thinking in a computer system, the chance of having a generic technique that can accurately infer semantic-rich analytic provenance information for a variety of analysis tasks is not high.

Instead, existing methods employ the following strategies:

1. Constrain the problem and analysis domain;

2. Aim for less semantically rich analytic provenance.

By limiting the choice of data and analysis/visualization, an inference algorithm has better chance to make the right guess of what analysis user is doing. However, even within a specific industry (such as finance), the types of data and analyses involved are still of very large amount. Also, being very limiting on the data and analysis will considerably constrain the system capability, having a negative impact on the sensemaking task. Therefore, this is likely to be used as a secondary approach, rather than the main one.

Given the difficult of inferring sensemaking information such as what analysis a user is doing, a few methods target less semantic rich but nonetheless potentially very useful analytic provenance information. One such example is "action chunking", i.e., identify a group of actions that are likely performed for the same sub-task, without knowing what the sub-task is. Such approaches apply heuristics to infer patterns from action logs based on repeated occurrence and proximity in data/visualization space or analysis time [5]. Such chunking information can be useful in many different ways. For example, the system can prompt user taking a note if such an action usually occurs within a specific sequence. Also, the grouping information can be used for aggregation when large amount of analytic provenance information is to be visualized.

2.3 Hybrid Capture

There is certainly space for the development of "hybrid" approaches, i.e., mixing the manual and automatic capture. The previous case of "action chunking" is such an example: the automatic capture of sub-task grouping can be improved with user input on whether a group of actions are a sub-task and what the task is. The improved algorithm can in turn help improve the manual capture by prompting users to take notes if such an action is expected within certain "action chunks" from previous experience. Similar approaches can be used to address other provenance capture challenges such as encouraging the recording of analysis rationale and process.

3 Sensemaking Provenance Visualization

Most existing provenance visualization methods focus on the action layer, which can be automatically captured and still offers certain level of semantic information. Often included are the user notes and links between provenance information within and across layers. Finally, each action or note is associated with its data, analysis, and/or visualization, which are usually included in the visualization, too. In most cases it is difficult to show all these information at once. Instead, existing methods often display selected provenance based on their design goals, with details on demand.

Node-link diagrams are a popular choice among methods that aim to show an overview of the sensemaking process through the sequence of user actions. In such methods, nodes represent system state and the edges represent actions that transit system from one state to another. While node-link diagrams can provide an overview of the sensemaking structure, they do not provide sufficient information for the understanding of the analysis process. Including related information, such as system state and user notes, can alleviate this problem. The most common approach is multiple-coordinated views that show the note and system information only for a selected step. This usually works well with many visual analytics systems, which already have view for each type of information; showing the analytic context essentially restores the system to a previous state. However, such setup still requires users to go through a process step by step, sometimes back and forth, to understand an analysis sequence, which places heavy cognitive work load on the user's memory. Methods such as GraphTrail [3] show multiple system states and the links between them at the same time. By allowing zoom and pan on the user interface, users can choose between overview of the analysis structure and details of individual system state.

Besides providing a general understand of the sensemaking process, analytic provenance can support more specific sensemaking tasks. One such task is (visual) narrative construction, which connects a series of findings discovered through the sensemaking processing to form a coherent "story". A narrative can include raw data, analysis results, visualization, and user notes. Narratives describes not only the sensemaking outcomes but also the process that leads to them, a useful feature for reporting and team collaboration. With analytic provenance, the efforts required for creating a narrative is considerably reduced. The DIVA system [14] allows interactive construction of narratives from captured analytic provenance and provides a summary of associated spatial and temporal uncertainty (Figure 2). SchemaLine [8] allow users to create hypotheses or narratives by grouping notes along the timeline (Figure 3). Analytic provenance has also been used to help users review their sensemaking

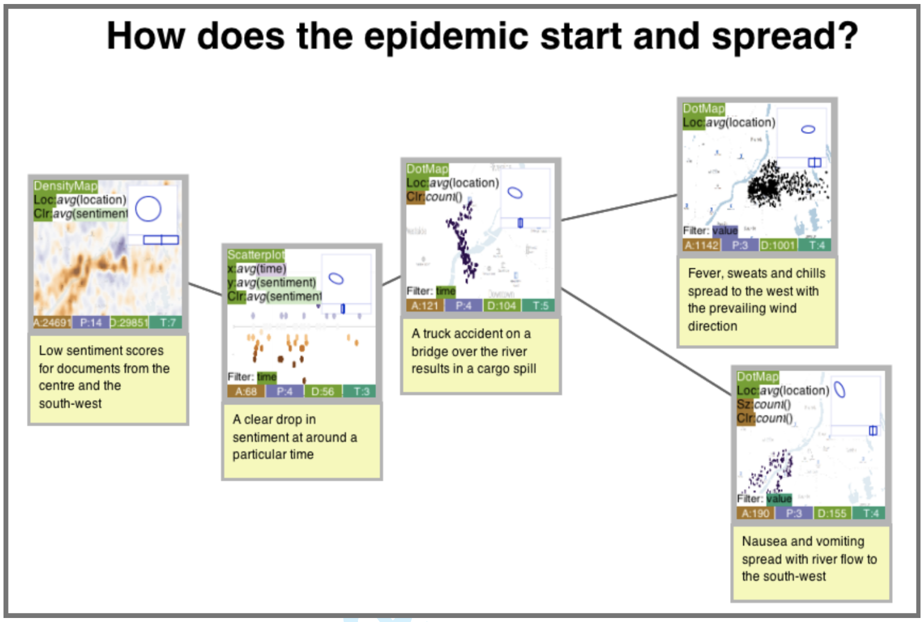

Figure 2: A narrative created in DIVA: each bookmark (box) is a saved visualization state (including the uncertainty information), together with the note (text at the bottom). Related bookmarks are linked together to form a narrative [14].

process and guide further exploration, which is particularly useful for analysis of complex dataset such as those with high dimensionality. Such methods [7,11] visualize the sensemaking space so user can easily see which part has been explored, e.g., which data dimension and which values within that dimension have been analyzed. Users can use this information to plan their further analysis and system can also use this information to suggest related but unexamined data.

4 Provenance for Collaborative Sensemaking

Part of our research agenda involves the role the provenance information offers for enhancing coordination during collaborative sensemaking. Collaboration is a process whereby multiple agents engage in joint activities aimed at advancing shared goals. It entails cooperation, coordination, communication and awareness. Awareness is essential for keeping track of others' activities as well as enabling a shared understanding of the work domain, i.e. 'a common situation picture'. Ideally, such awareness promotes contributions which are aligned in ways which are fluent, seamless, unpremeditated and inconspicuous.

The need for visibility is demonstrated to some extent through the use of "coordination artefacts" that collaborators sometimes use, such as written plans,

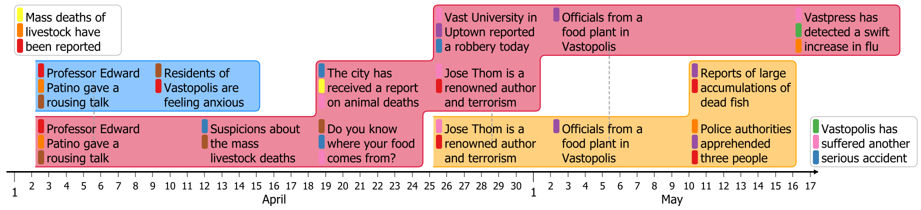

Figure 3: In SchemaLine, each piece of text is an analyst note, positioned along the time axis according to the temporal information in the associated data. Users can link related notes to form a "schema", which can be either a hypothesis during the early stage of analysis or a narrative to present the final findings. There are three schemata in this example represented as differently colored rectilinear paths [8].

procedures, timetables, schedules, checklists and other mechanisms which offer cues about intentions and action. By providing a trace of activities of collaborators who may be acting at a distance and asynchronously, provenance information has the potential to play an important role in providing cues for collaboration. Seeing the record of the actions of others allows the inference of their intent that may not be present in their results alone. As such, one of our research issues is the coordination—or handover—of provenance between collaborators.

In addition to supporting collaboration around common goals, provenance information can also provide a basis for sharing best practice. What counts as best practice may not be immediately evident and may need to be identified over time and in relation to pre-defined success measures or indicators. Nevertheless, capturing the way that tasks have been tackled through provenance provides an opportunity for reconstructing successful approaches and identifying their significant features after the fact. This could then provide a basis of training and optimizing processes. Conversely, provenance may also provide case studies for failed processes for training. However, using failed processes without damaging the reputation of the involved users is an open research problem.

4.1 Privacy

A key issue to be considered when designing any system which involves the recording and retrieval of people's actions is that of privacy. It may be unethical or even illegal to record all of a user's actions performed on a system without their prior permission to do so. In the design of systems which record provenance, designers need to consider exactly what data is recorded, what that data will be used for, and by whom. Depending on context, it may or may not be appropriate to design a system which is sufficiently 'socially translucent' such that people can be held accountable for their actions [4]; privacy may require that there are contexts in which they should not.

Furthermore, the way in which provenance is captured, and the level of accountability that can be embedded in data captured by a system, may affect the way people use it. For instance, when data is aggregated and individual users are unable to be identified, they may be less reluctant to explore or experiment, as a level of plausible deniability preserved. Conversely, when a system is able to record individual user identities and their actions, people may be less inclined to perform certain actions in fear of recrimination from their superiors.

One potential solution to this is the preservation of anonymity. But this poses a further issue: when users collaborate, it can be important for them to be able to identify other members of the team and their contribution. In order to remedy this, a system may incorporate some level of internal/external split, meaning that some information can only be held internally within a group. Additionally, there may be different levels of privilege within a system according to a hierarchical structure, with different levels of access according to a user's level within that structure. However, the acceptance of this from a cultural perspective must be taken into consideration. Cultures where there is a higher power distance and more well defined hierarchy will find such solutions more acceptable than those with a more equal distribution of power.

4.2 Handover of Provenance

Essential to collaborative work in many disciplines is a well established handover process. Handover refers to the reporting of the state of the world at a given point in reference to work being carried out. In fields such as medicine and nursing, there are well established practices and protocols for carrying out handover activities [9]. A great deal of work has been carried out in understanding handover and provenance in this medical domain. Further work should be carried out to extend this understanding into fields such as intelligence analysis or software development.

Performing handover activities could involve collapsing a potentially complex state-space into a means where it can be effectively and efficiently articulated to a receiving agent in order to continue work. We see this in patient transfer and other parts of the medical domain, where there is an established practice of handover essential for the continuing care of patients [9]. Nurses and medical professionals have developed an efficient protocol and vocabulary for transferring care of patients; a general vocabulary of this nature is an open problem. In the development of collaborative systems, we argue that the representation of provenance information in order to handover work holds great importance. By exploring existing work and carrying out further studies 'in the wild', we can establish an understanding of the level of abstraction being used, and how a state-space is being represented in current work setting, providing vital insights to the design and development of collaborative systems.

5 Uncertainty and Trust

Underlying the challenges faced by visualizing provenance and understanding its use in collaboration are questions about the validity of the process and its record. Original data may be of low quality, depictions and interactions with the data may exacerbate uncertainty, leading to a lack of trust (or over-trust) of the result [10]. These two issues—uncertainty and trust—present significant challenges in the successful use of analytic provenance in sensemaking.

5.1 Uncertainty

Uncertainty, in our context, are variations from the stated value introduced to our data before or during its analysis. Before analysis, uncertainty stems from lack of precision in measurement, inconsistencies in recorded results, or missing values. During the analysis, tools in the workflow can modify or introduce uncertainty—sampling and aggregation, such as that done to ensure privacy, transforms even certain values into a representational result with some variance from the population. The result of any individual visualization is thus uncertain, even if it is presented in a manner that hides this variation.

From our workshop discussion, there are three main challenges driven by the uncertainty in the analytics process. First, it is unclear from a general standpoint how to characterize uncertainty—what are the appropriate metrics for different types/sources of uncertainty, and how do they appropriately propagate through workflows? [10] Error analysis is well studied for arithmetic operations, but how do they combine under sampling, aggregation, or other transformation? How do we quantify and propagate uncertainty due to multiple witness statements, intelligence reports, or other non-quantitative measures? There is a research opportunity to characterize a reusable typology of uncertainty factors with known propagation methods. This typology will likely be built from domain specific examples of uncertainty first before a more general model is known.

The second, related challenge relates to using the uncertainty to guide insight discovery. Even if the uncertainty in the process is understood, it is unclear how to model what the user currently knows about the data (and its certainty) or the extent of the analysis space covered. Metrics about the exploration process can assist [6], and the methods alluded to in the previous sections can partially address this challenge. There is a research challenge to integrate a model of uncertainty into these recommendation and insight modeling systems. Especially challenging is modeling and highlighting the unknowns—what is the uncertainty hiding in the data, or what are the range of valid results.

Both of the previous challenges require an understanding of how the uncertainty affects the user's understanding and their sensemaking. Thus, how to synthesis understandable uncertainty that fits the user's model of uncertainty is our final research challenge. Sensemaking under uncertainty needs to be studied to characterize and mitigate misunderstandings that occur due to this inherit lack of information. While every sensemaking tasks begins with lack of knowledge of the final result(s), this "unknowledge" is different from sensemaking under uncertainty, where the process itself cannot be given full trust. Understanding the best practices for mitigating uncertainty in the process will assist users make decisions under uncertainty.

5.2 Trust

Even in an certain sensemaking process, levels of trust in the results may vary; uncertainty makes trust decisions more fraught. In analytics for sensemaking, trust appears in three contexts: Trust in the data, trust in the process, and trust in the result. Each presents challenges for research.

Trust in the data is an issue of data quality. Uncertainties introduced in the measurement, storage, and access of data all affect the trust in its validity. Provenance of the the data sourcing and workflow can be used as part of trust decisions regarding that data; as a computational artifact, data quality provenance can also be part of synthetic trust models [1, 13]. The research problems here are both on the representation of the data's quality (what are the appropriate metrics? How do these interact with our uncertainty propagation models as part of the workflow?) and on its communication to the user (how to indicate when a user is making risky inferences from data under low quality conditions? How do we depict the consequence of different quality representations in terms of workflow computational usage or result fidelity?).

A user's trust in the analytical process, while related to uncertainty, incorporates other measures—the user's trust in the data, their believed expertise of the material, and the cognitive biases they bring to the analysis. How can a computer synthesize a model of trust built form these factors? Venters' et al. [13] and Sacha [10] suggest tracking the provenance of the data and the analytical process to measure a user's trust in the process—tighter exploration loops suggest confidence whereas scattered exploration suggest distrust of the process An open research challenge is to formally measure and quantify a trust inference model from given user explorations. While examples have been gathered, such as classroom visualization usage [12], but more work is needed to generalize. It is also an open question of how to detect and communicate biases in the analytic process; inferring when a user is not exploring potentially fruitful avenues due to unconscious inattention is vital in robust recommendation systems.

Trust in the result is tied to their confidence in the process and the original data. Previously, we spoke of a model for the user's inference and thus confidence of the process and result; in concert with that model would be one for measuring the risk associated in using the result. If uncertainty cannot be eliminated, it could be mitigated if appropriate measures of risk could be devised [13]. Determining appropriate risk models is an open research problem, and tying the risk to the uncertainty/trust is also an open challenge.

6 Conclusions & Research Agenda

Visual analytics can be improved via a better understanding of the behavior during the analytic process in support of sensemaking—provenance can be used for reporting and training, can facilitate collaboration, and help us understand what we can trust from our possibly uncertain data. We have presented several challenges raised from these topics during our IEEE VIS 2014 workshop as part of a research agenda for the community. Taken together, they form a four part research agenda:

- **Enhance provenance capture** to better support more accurate and higher level inference from analytic provenance. These may be manual, automatic, or hybrid, but such inference can assist in understanding the provenance process for better prediction, process correction, and decision making.
- **Develop and validate sensemaking support visualizations** from analytic provenance. Current research has only scratched the surface of the semantically rich space of information present in the provenance; to support the enhanced provenance capture recommended above, additional visual presentations are needed.
- **Investigate privacy-aware methods to utilize collaborative provenance** that provide the appropriate level of detail depending on the sensemaking task and the role of the user. Proper collaboration will also require general summarization and provenance handover between collaborators.
- **Extend error propagation through provenance pipelines to wider types of uncertainty** via better typologies and studies of sensemaking risks under uncertainty. This agenda is synergistic with enhance provenance capture—better intent inference can be used to build model of trust in the sensemaking, whereas improved uncertainty models can correct over trusting inference models.

Systems and practices for supporting sensemaking are a vital part of the larger visual analytics context. We see that in the future, as visual analytics broadens its reach, better support for sensemaking will require solving these and other analytic provenance challenges.

References

[1] C. Bors, T. Gschwandtner, S. Miksch, and J. Gärtner. Qualitytrails: Data quality provenance as a basis for sensemaking. In *Proceedings of the International Workshop on Analytic Provenance for Sensemaking*, 2014.

[2] W. Dou, D. H. Jeong, F. Stukes, W. Ribarsky, H. R. Lipford, and R. Chang. Recovering Reasoning Processes from User Interactions. *IEEE Computer Graphics and Applications*, 29(3):52–61, May 2009.

[3] C. Dunne, N. H. Riche, B. Lee, R. A. Metoyer, and G. G. Robertson. GraphTrail: Analyzing Large Multivariate and Heterogeneous Networks while Supporting Exploration History. In *ACM Conference on Human Factors in Computing Systems*, pages 1663–1672, 2012.

[4] T. Erickson and W. A. Kellogg. Social translucence: An approach to designing systems that support social processes. *ACM Transaction on Computer-Human Interaction*, 7(1):59–83, Mar. 2000.

[5] D. Gotz and M. X. Zhou. Characterizing users' visual analytic activity for insight provenance. *Information Visualization*, 8(1):42–55, 2009.

[6] T. J. Jankun-Kelly. Using Visualization Process Graphs to Improve Visualization Exploration. In J. Freire, D. Koop, and L. Moreau, editors, *Provenance and Annotation of Data and Processes*, volume 5272 of *Lecture Notes in Computer Science*, pages 78–91. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[7] A. Lunzer. Lightweight provenance-driven exploration. In *Proceedings of the International Workshop on Analytic Provenance for Sensemaking*, 2014.

[8] P. H. Nguyen, K. Xu, R. Walker, and B. W. Wong. SchemaLine: Timeline visualization for sensemaking. In *2014 18th International Conference on Information Visualisation (IV)*, pages 225–233, July 2014.

[9] R. P. Variations and commonalities in processes of collaboration: the need for studies of multiple settings. *Computer Supported Cooperative Work: CSCW: An International Journal*, 20(1-2):37–59, 2011.

[10] D. Sacha, H. Senaratne, B. C. Kwon, and D. A. Keim. Uncertainty propagation and trust building in visual analytics. In *Proceedings of the International Workshop on Analytic Provenance for Sensemaking*, 2014.

[11] A. Sarvghad and M. Tory. Exploiting history to reduce interaction costs in collaborative analysis. In *Proceedings of the International Workshop on Analytic Provenance for Sensemaking*, 2014.

[12] C. T. Silva, E. Anderson, E. Santos, and J. Freire. Using vistrails and provenance for teaching scientific visualization. *Computer Graphics Forum*, pages 75–84, 2011.

[13] C. C. Venters, J. Austin, C. E. Dibsdale, V. Dimitrova, K. Djemame, M. Fletcher, S. Fores, S. Hobson, L. Lau, J. McAvoy, A. Marshall, N. T. Townend, V. Viduto, D. E. Webster, and J. Xu. To trust or not to trust? developign trusted digital spaces through timely reliable and personalized provenance. In *Proceedings of the International Workshop on Analytic Provenance for Sensemaking*, 2014.

[14] R. Walker, A. Slingsby, J. Dykes, K. Xu, J. Wood, P. H. Nguyen, D. Stephens, B. L. W. Wong, and Y. Zheng. An extensible framework for provenance in human terrain visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2139–2148, Dec. 2013.